THE USE OF ARTIFICIAL INTELLIGENCE AND BIG DATA IN THE SAFETY EVALUATION OF FOOD ADDITIVES

by
Yuqi Fu

A thesis is submitted to Johns Hopkins University in conformity with the requirements for the degree of Master of Science

Baltimore, Maryland
May 2021

Abstract

Environmental contaminants, naturally occurring toxicants, pesticide residues, and food additives are the four chemical-associated of six categories of food safety established by the Food and Drug Administration. The direct food additives, which are intentionally added to the food, are the focus of this research, and the indirect food additives, such as pesticides, natural toxicants and environmental residues will also be discussed. This study is attempting to investigate how artificial intelligence and big data could benefit the evaluation of food additives.

Automated Read-Across technology, i.e., the read-across-based structure activity relationships (RASAR), are utilized as an example to compare with traditional animal testing methods to assess their utility for providing accurate enough estimates of chemical toxicities for food-relevant substances. The comparison shall be conducted using Underwriters Laboratories (UL) Cheminformatics Tool Kit and then following up with descriptive statistics manipulated by Microsoft Excel and validation datasets retrieved from other sources such as ECHA (the European Chemicals Agency), EPA (the U. S. Environmental Protection Agency), OSHA (the Occupational Safety and Health Administration), EFSA (the European Food Safety Authority), and other literature.

After statistical analysis, main findings were listed below. It was rare to have two manual curation categories for one chemical. Generally, one chemical corresponds to one manual curation category. There were more direct food additives and indirect food additives in the training data. In this chemical list, there were more non-toxicants than toxicants, which was expected for food-related substances. More results were founded at very strong and strong confidence level. 83% of the Read-Across results selected for validation process match with the toxicological assessment results from other sources and literature.

Primary Reader and Advisor: Thomas Hartung
Secondary Reader: Keeve Nachman

Contents

List of Tables

List of Figures

1. Introduction

1.1 Food safety and food additives

As human beings, we rely on food to sustain our lives. It is essential to eat sufficiently and healthily. However, the problem is that not everything that can be eaten is safe. Improper production or handling of food could lead to severe consequences. As part of the progress of modern technologies, various kinds of food preservation such as canning food appear, and different types of processing methods are developed accordingly, which makes the public pay more attention to food safety issues and food and water-borne diseases (Mitchell et al., 2007). Food safety mainly refers to how industrial and regulatory organizations manufacture, store and deal food in order to prevent foodborne disease. In China, there is a proverb that says, "*illness finds its way in by the mouth*", which implies the tight connection between food and human health and the importance of food safety in human life. Every food that enters the market should be tested strictly and have a label of ingredients it includes. This rule shows that what every country's government does regarding food safety is a significant topic, and in 1962, John F. Kennedy declared consumer safety as a fundamental right (John F. Kennedy, 1962). Food safety has two fundamental aspects, i.e., microbiological and chemical safety. More specifically, the Food and Drug Administration established six categories of food safety hazard and national issues, which include microbial contamination, nutritional problems, environmental contaminants, naturally occurring toxicants, pesticides residues, and food additives. Food additives, and in part pesticides/residues and other toxicants, shall be addressed in this research. Food safety regimens need to ensure every additive added to the food is safe for human beings to eat, and therefore there are several organizations established to test and evaluate the ingredients to see if there are any severe negative effects or toxicities. The testing methods include traditional methods like animal testing and newly developed methods such as *in silico* testing, which make sure the food can be consumed without concerns.

During the process of dealing, manufacturing and storing food, the development and utilization of food additives are common. If you pay attention to the ingredients of food, like some snacks, you could find many types of food additives just in one pack of snacks. The function of food additives includes preservation, flavor, color, sugar substitutes, enzymes,

acidulates and so on. FDA defines that food additives are not limited to things that are added to the food, which are named direct food additives, but also things that can contact and contaminate the food during the process of dealing, manufacturing and storing. Pesticides/residues and packaging material are considered as indirect food additives (Meigs et al., 2018). Albeit some manufacturers tried their best to replace food additives by some natural materials, the trends of using food additives still increases globally, especially for artificial sweeteners (Mordor Intelligence, 2019). According to a report, the year-on-year growth rate of seasoned and flavoring salt increased 6.88% per year from 2016 to 2018 globally (Mordor Intelligence, 2019). As the growing demands of food appearance, texture, flavor and taste, the usage of different types of food additives will still be rising in the next years, it is extremely important to come up with more convenient and effective methods to cope with those thousands of chemicals as food additives for the sake of saving time and increasing accuracy.

1.2 Regulation of food additives

Regulations of food additives throughout the world have several similarities and dissimilarities. For example, the full extent of engineered nanomaterials that entered the U.S. cannot be known by FDA due to it has been considered as Generally Recognized as Safe (GRAS) substances, which are the substances that been considered as safe by experts and could be exempted from the usual Federal Food, Drug, and Cosmetic Act (FFDCA). However, in Canada and the European Union, all food contains engineered nanomaterials should be submitted to regulators before entering the market (GAO, 2010). The global regulation of food additives is mainly led by a few safety assessment programs, which are the European Union Scientific Committee on Food, the European Food Safety Authority (EFSA) and the United States Food and Drug Administration (FDA) (Mitchell, 2014).

The US food additive regulation has some particularities for food additives or ingredients, and the most unique thing in the US regulation is that they include the indirect food additives and exempt those generally recognized as safe by qualified experts as the substance had been sufficiently showed as safe from premarket approval requirements (Mitchell, 2014). The FDA needs to provide a regulation including any specifications and limitations to show the intended use of the additive is safe (Code of Federal Regulations, 2013). The U.S. process contains estimation of dietary consumption, assessment of likely toxicity and a risk management decision regarding safety (Mitchell, 2014). This process sometimes will also take nutritional habits and

sensitive subpopulations into account (Mitchell, 2014). The FDA's assessment of new chemicals ought to be conducted based on sufficient available toxicological data and proper estimated dietary exposure. The process of addressing sufficient data will involve the consideration of testing according to the FDA Toxicological Principles for the Safety Assessment of Food Ingredients Redbook (FDA, 2000). Studies need to be reviewed based on the Redbook and conducted in compliance with good laboratory practice regulations (Code of Federal Regulations, 2013). The process of determining proper estimated dietary exposures is followed by the actual procedure of conducting the test and determining the point of departure, such as any adverse effects, no adverse effect levels, low effect levels, and other indications of toxicological concern.

The EU food additives regulation does have several differences from the US food additive regulation such as instead of self-regulation by industry a mandatory premarket authorization. Also, cyclic reconsideration of the safety of additives is the mandatory. The EU treatment process of processing aids, the substances which add to processed food and remain in the finished product, but it is not required to provide as an ingredient for consumers by law (International Food Information Council, 2014), does not differ a lot from the regulation of GRAS except it is more limited in scope. JECFA for other countries also follows the same general approach for assessment.

Back to GRAS, even though the FDA states clearly that both food additives and GRAS ingredients require the same strength of evidence of safety, "*currently, companies may determine substances are GRAS without FDA's approval or knowledge. However, a few substances previously considered GRAS have later been banned; and concerns have been raised about the safety of other GRAS substances, including those containing engineered nanomaterials, materials manufactured at a tiny scale to take advantage of novel properties*." (GAO, 2010). We can see that GRAS is not completely equal to safe and needs to be evaluated regularly. Hartung (2016), suggested the following approach to GRAS evaluation: 1) determining the GRAS eligibility of the substance, 2) collecting all available information on the substance separately for every information need, 3) considering a Threshold of Toxicological Concern (TTC) approach (Hartung, 2017), 4) developing a test strategy, 5) carrying out the test strategy and the respective risk assessment, 6) considering mixture effects, sensitive subpopulation and extreme use scenario, 7) evaluating metabolites, degradation products and impurities, and 8) documenting the

process and share the results with FDA, and the general public (Hartung, 2018). As time goes by, an increasing number of new technologies will appear and carry out more new ways to replace previous testing methods and to cope with chemicals which cannot be tested readily before. In that case, GRAS should also be retesting regularly and previously existing methods for GRAS ought to be revised, or else GRAS may lead to potential public health concerns.

1.3 Testing methods

The testing methods used for evaluating food additives include both traditional methods, such as animal testing, in vitro testing, and newly developed computational methods. Animal testing uses experimental animals to determine the toxicity of the substances. *In vitro* testing is conducted using components of an organism that have been isolated from their usual biological surroundings, such as microorganisms, cells, or biological molecules. *In vivo* testing studies the effects on various biological entities such as whole, living organisms, usually animals, sometimes including humans, and plants, as opposed to a tissue extract or dead organism. *In silico* assessment is part of the non-testing methods that is conducted through some computational programs and relies increasingly on big data, i.e., large, diverse sets of information. Currently, alternatives to animal tests follow the philosophy known as the 3Rs – replacement, reduction, and refinement. This concept attempts to reduce or exclude animals from the tests, or alleviate the pain and distress posed on animals (MacArthur, 2017).

Generally, animal testing is most common among the testing methods for food additives in previous years. However, these tests can be too costly, time-consuming and sometimes lead to misleading results. Due to the high demand for testing data, there could be many animal tests conducted to test only one specific chemical with reliable results. Individual tests can also require a large population of test animals to diminish possible bias. We can easily imagine how costly animal tests can be, amounting to several million dollars for the full evaluation of a single substance. During the process of food additives risk assessment, researchers will attempt to determine and establish the NOAEL value. This determination requires the researchers to repeat the test on many experimental animals for several different doses, or else the most appropriate NOAEL value cannot be established and accepted. Some animal tests can lead to ambiguous results, which requires follow-up testing to address this problem. In that case, the time invested in this test can be much longer. However, the problem is that even tests without ambiguous results can cost researchers too much time, this is the problem of prolonged testing periods, e.g.,

it can take easily five years to obtain the result of a cancer bioassay, again slows down the evaluation process, which by no means can be good news give the growing number of untested chemicals. The other important issue is that we need to extrapolate animal testing results to human beings. Due to some ethical concerns, it is hard and often improper to conduct human testing for food additives. Hence, researchers utilize experimental animals as surrogates for humans and extrapolate the results, which cause hidden problems. For example, the exposure routes include ingestion, inhalation, and dermal contact, and ingestion and inhalation routes could be hard to draw analogies. Specifically, animals and human beings have different digestive tracts and respiratory systems, and chemicals can be processed by different mechanisms in the body, which make the extrapolation difficult and not that accurate. For example, experimental rodent animal's anatomy and physiology differs from humans. Other weaknesses for extrapolating animals testing results to humans include humans have more comprehensive coverage of life stages, more systematic evaluation of pharmacokinetics of fetuses or young animals, more focused evaluation of structural and functional toxicity in the young before and immediately after weaning, more organ-specific evaluations include immune, nervous, and endocrine systems, and the lack of a systematic approach for carcinogen testing. Altogether, it is easy to overestimate the accuracy and effectiveness of animal testing results. The in vitro testing also has similar disadvantages such as being hard to extrapolate the results back to an intact organism. The *in vivo* testing has the disadvantage of high costs too.

The ToxCast high-throughput screening program, a computational approach based mainly on *in vitro* high-throughput screening assays, provides us with a quick insight into the biochemical endpoints, cellular processes, and phenotypes for sizeable fraction of 8,659 food-relevant chemicals (Karmaus et al., 2016). Karmaus divided chemicals into three categories, which are direct food additives, food contact substances, and pesticides, and obtained the databases from several FDA resources. For direct food additives, chemicals naturally occurring in foods also considered as direct food additives in Karmaus work (Karmaus et al., 2016). The data in ToxCast are open source (website).

Compared to animal testing, the alternative of Read-Across, which is the currently most frequently used non-animal alternative approach, that predicts hazard from chemical analogs with known hazard data (Hartung, 2016), can do a better job. It could save time and money to large extent and when automated possibly be more accurate as shown by Luechtefeld et al.

(Luechtefeld et al., 2018). For very large datasets, it is more convenient to input the data in a computational program to see the patterns, trends, and associations. According to Hartung, "*Big data can nurture the ugly duckling [of read-across] to becoming a beautiful swan.*" (Hartung, 2016). Read-Across has been combined with quantitative structure - activity relationships (QSAR), which is a classification model, called the RASAR approach (read-across-based structure activity relationships). The model is used for predicting simple effects such as skin sensitization, eye irritation, and other side effects based on big data. The database includes approximately 10 million chemicals and enables the identification of highly similar compounds. "*This allows inference of respective properties in a process that is called read-across.*" (Patlewicz et al., 2014). In conclusion, besides being cheap, time-saving and accurate, the dataset can also assess the toxicities and properties based on the similarities, determine the frequency of the hazards, evaluate the quality of previously conducted animal testing, and monitor the REACH registration (Hartung, 2016).

Food additive is a national food safety concern, but the utilization of artificial intelligence and big data could assist to cope with untested substances and impurities in drugs and food. These analyses would be assessing the performance of the RASAR for food-relevant substances, identifying selected food-related substances that are out-standing as toxic or discrepant between actual and predicted properties, attempting an external validation by identifying chemicals with test results outside the Read-Across technology.

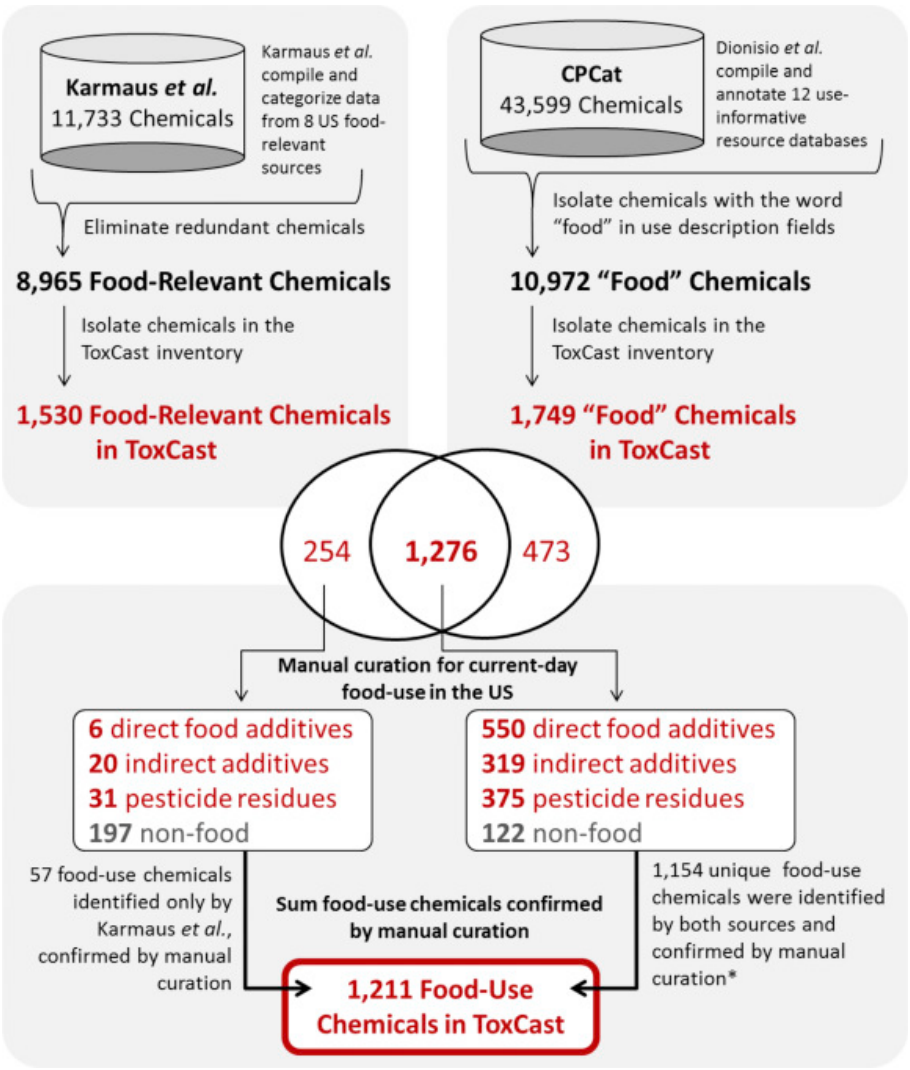
2. Materials and Methods

2.1 Accessible databases

The chemical list used for further statistical analysis was generated from the combination of three datasets. Two of the three datasets are "SuppFile2_SOMbins_SMILES" and "1,749 CPCat ToxCast details" retrieved from Dr. Karmaus' works published in 2016 and 2017, and the other one is "Read-Across results" provided by Dr. Luechetfeld after running Dr. Karmaus' chemicals in "SuppFile2_SOMbins_SMILES" through the Read-Across program (UL Cheminfomatics Toolkit 2.0, ULCT).

Dr. Karmaus' 2016 work presents the complication of a large inventory of food-used chemicals including direct, indirect/food contact, and pesticide residue chemicals. Analyses in their work go on to limit the inventory to only those compounds that overlap with ToxCast (Karmaus et al., 2016). The 2017 work provided comprehensive manual curation of the food-used chemicals in ToxCast, which was conducted to refine the inventory to reflect food additives used in current days. The 2017 work excluded some chemicals from the 2016 work that had once been registered as food additives but in more recent years had been removed from such use. The 2017's curation of food-relevant chemicals in ToxCast is only a curation of the subset that was overlapping with ToxCast, and curation was based on current use in the United States (Karmaus et al., 2017).

Figure 2.1 Workflow for the identification of US-relevant ToxCast food-use chemicals



Initially, Dr. Karmaus came up with 11,733 chemicals from 8 US food-relevant sources, which are Everything Added to Food in the US (EAFUS), Generally Recognized as Safe (GRAS) Notice Inventory, Select Committee on GRAS Substance Database (SCOGS), List of Indirect Additives Used in Food Contact Substances, Inventory Effective Food Contact Substances, Threshold of Regulation (TOR) Exemptions, the Flavor and Extract Manufacturers Association GRAS Inventory (FEMA), and the Aland Wood Pesticides database, at the very beginning. After eliminating redundant chemicals, she got 8,965 distinct chemicals. Then, by distinguishing chemicals with discrete structures, 8,965 chemicals narrowed down to 4,729 chemicals, which all of them have clear structures (Karmaus et al., 2016). These 4,729 chemicals not only underwent further analysis by Dr. Karmaus to get the manual curation information, such as direct food additives, indirect food additives, pesticides/residues, and non-food, but were also

run through the Read-Across technology (UL Chemifomatics Toolkit 2.0, ULCT) by Dr. Luechetfeld to provide the prediction of 9 different health and environmental endpoints, which are acute toxicity, acute dermal irritation, acute dermal toxicity, acute aquatic toxicity, acute inhalation toxicity, chronic aquatic toxicity, eye irritation, mutagenicity, and skin sensitization. In Dr. Karmaus' further analysis, 4,729 chemicals were narrowed to 1,530 food-relevant chemicals, which have ToxCast details. Dr. Karmaus also used 43,599 chemicals from 12 use-informative resource databases, which are Aggregated Computational Toxicological Resources (ACToR) data sets and lists, ACToR UseBB, Design for the Environment (DfE), Dow Chemical Company (Dow), Drug Bank, U.S. EPA 2006 Inventory Update Reporting (IUR) Modifications Rule and the 2012 Chemical Data Reporting (CDR) Rule, Swedish Chemicals Agency (Keml) National Industrial Chemicals Notification and Assessment Scheme (NICNAS), Retail Product Categories (RPC) database, Substances in Preparation in Nordic Countries (SPIN) database, and Human Toxome Project (HTP) (Dionisio et al., 2015),based on Dr. Dionisio's work. After isolating chemicals with the word "food" in use description fields, 43,599 chemicals narrowed down to 10,972 "food" chemicals. These 10,972 "food" chemicals then narrowed to 1,749 "food" chemicals in ToxCast by isolating chemicals in the ToxCast inventory. Then, the list of 1,530 chemical was compared to the CPCat (the EPA Chemical/Product Categories database) to see if it could identify any use information at all, and only 1,530 food-relevant chemicals in ToxCast and 1,749 "Food" chemicals in ToxCast underwent manual curation (Karmaus et al., 2017). This manual review provided the CASRNs (chemical abstract services registration numbers) and manual curation categorizations for corresponding chemicals. In our Read-Across analysis, 4,729 chemicals have been inputted because these chemicals had clear structures and SMILES (simplified molecular input line entry system codes), which is the required information to run a chemical analysis in the ULCT. In Read-Across, each chemical in 4,729 chemicals underwent 9 analyses based on 9 different health and environmental endpoints, which are acute oral toxicity, acute dermal irritation, acute dermal toxicity, chronic aquatic toxicity, eye irritation, acute inhalation toxicity, acute aquatic toxicity, mutagenicity, and skin sensitization. The results provided 38,520 prediction results for 4,729 chemicals. Prediction value 0.5 was taken as the threshold for calling a chemical positive for a given hazard, i.e., prediction values equal to or above 0.5 would be considered as adverse for the health endpoint. Prediction values below 0.5 would be considered as negative, which means the tested chemical is predicted not to

lead to the tested adverse health endpoint. The reliability string, which is the confidence level of the prediction value, for each prediction value was obtained as well. Reliability string very strong and strong means high confidence level of the prediction value. Reliability string moderate and low means low confidence level of the prediction value. Reliability string "in training" means the chemical had data used for training the algorithm, and these results were part of the training dataset to build the prediction model. Reliability string "exclude" means the structure for the tested chemical does not fit ULCT rules.

2.2 Statistics

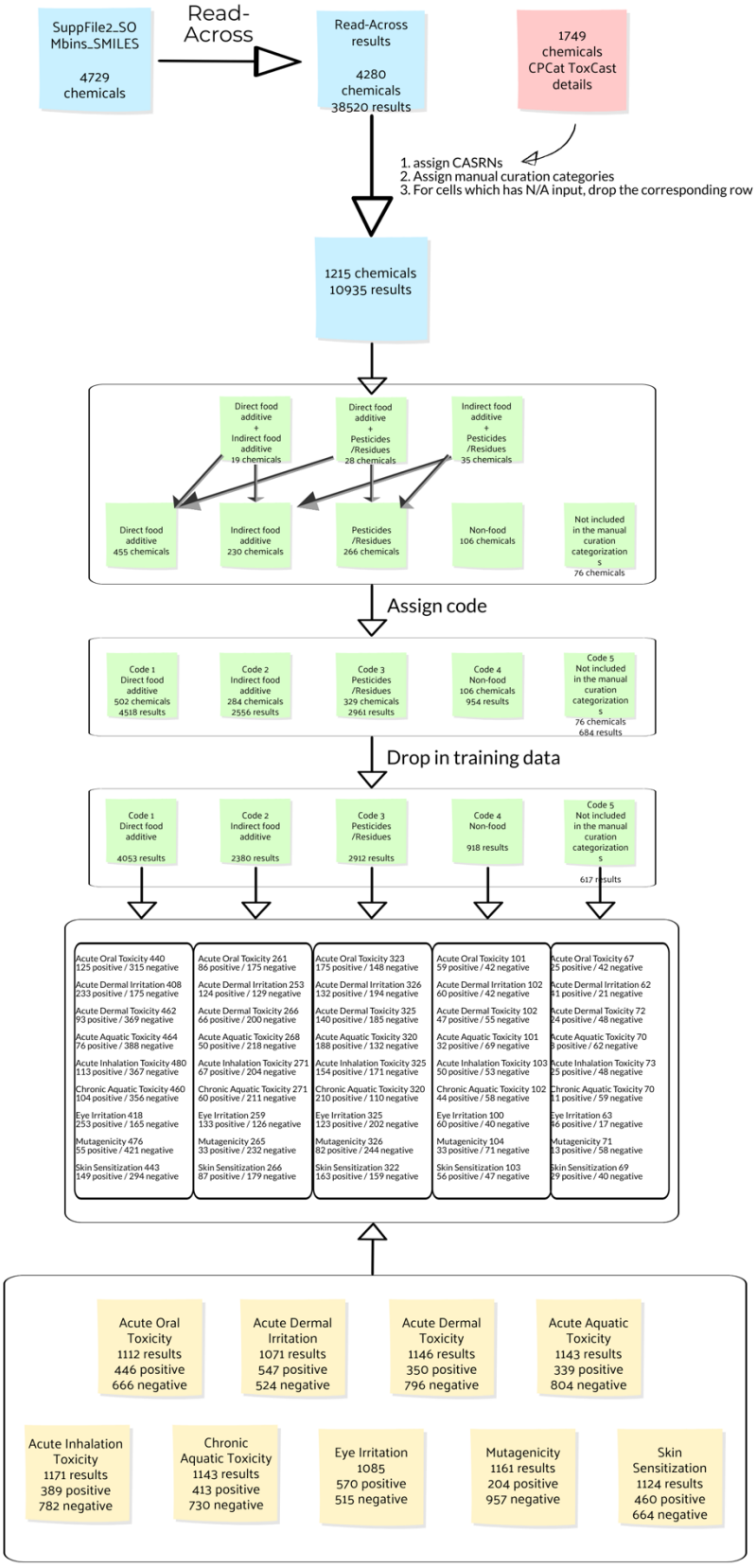
In order to get the chemicals, which could be used for statistical analysis, the first step is assigning CASRNs for every chemical in "Read-Across results". Using the VLOOKUP function in Microsoft excel and the SMILES identifiers information, CASRNs could be assigned to the corresponding chemicals. 38,520 results had been obtained for 4,729 chemicals. Then, the next step is assigning manual curation categorization for those 38,520 results based on CASRNs. The tool employed is the VLOOKUP function in Microsoft excel. After assigning manual curation, N/A results appeared in manual curation categorization column. The appearance of N/A indicated this kind of chemicals had not run through manual curation review in Dr. Karmaus' research, which means currently we are unable to provide the manual curation categorizations for those chemicals based on the information we have. Hence, the sort and filter function in Microsoft excel has been used to exclude those chemicals and those corresponding prediction results. Finally, 1,215 chemicals which have SMILES identifiers, CASRNs, predictions of health and environmental endpoints, confidence levels of predictions, and manual curation categorizations came up, and these 1,215 chemicals are the chemical list, which was used for the following descriptive statistical analysis. Corresponding to these 1,215 chemicals, there were 10,935 results with prediction value and confidence level.

These 1,215 chemicals could be divided into 8 manual curation categories. Those 8 manual curation categories are direct food additive, direct food additive + indirect food additive, direct food additive + pesticide/residue, indirect food additive, indirect food additive + pesticide/residue, non-food, pesticide/residue, and not included in manual curation. Code numbers had been assigned to each manual curation. 8 manual curations had been replaced by their corresponding code in the excel sheet. By using Microsoft excel SUMPRODUCT function,

the frequency of appearance of each manual curation in the 1,215 chemicals list can be calculated.

By using the filter and sort function in Microsoft Excel, results for different manual curation categories could be extracted based on key words. Because there are 8 manual curation categorizations, 8 separate sheets have been created to document corresponding results. For example, in the “Direct food additive” sheet, only results which labeled as direct food additive in the manual curation categorization column have been included. In each manual curation categorization sheet, the results could be further divided into subgroups by 9 health endpoints, and still use the filter and sort function in Microsoft Excel to create 9 separate sheets. For example, in “1AOT” sheets, only the results with direct food additive manual curation and acute oral toxicity health endpoint had been included. “1AOT” could still be divided into subgroup based on the confidence levels. The “reliability_strings” column are the confidence levels of each prediction. Using Microsoft Excel SUMPRODUCT function, the frequency of appearance of each health endpoint and confidence level could be counted in each sheet.

Figure 2.2 Workflow for the 1,215-chemical list


SuppFile2_SOMbins_SMILES
4729 chemicals
Read-Across
Read-Across results
4280 chemicals
38520 results
1749 chemicals CPCat ToxCast details
1. assign CASRNs
2. Assign manual curation categories
3. For cells which has N/A input, drop the corresponding row
1215 chemicals
10935 results
Direct food additive + Indirect food additive
19 chemicals
Direct food additive + Pesticides /Residues
28 chemicals
Indirect food additive + Pesticides /Residues
35 chemicals
Direct food additive
455 chemicals
Indirect food additive
230 chemicals
Pesticides /Residues
266 chemicals
Non-food
106 chemicals
Not included in the manual curation categorizations
76 chemicals
Assign code
Code 1 Direct food additive
502 chemicals
4518 results
Code 2 Indirect food additive
284 chemicals
2556 results
Code 3 Pesticides /Residues
329 chemicals
2961 results
Code 4 Non-food
106 chemicals
954 results
Code 5 Not included in the manual curation categorizations
76 chemicals
684 results
Drop in training data
Code 1 Direct food additive
4053 results
Code 2 Indirect food additive
2380 results
Code 3 Pesticides /Residues
2912 results
Code 4 Non-food
918 results
Code 5 Not included in the manual curation categorizations
617 results
Acute Oral Toxicity 440
125 positive / 315 negative
Acute Dermal Irritation 408
233 positive / 175 negative
Acute Dermal Toxicity 462
93 positive / 369 negative
Acute Aquatic Toxicity 464
76 positive / 388 negative
Acute Inhalation Toxicity 480
113 positive / 367 negative
Chronic Aquatic Toxicity 460
104 positive / 356 negative
Eye Irritation 418
253 positive / 165 negative
Mutagenicity 476
55 positive / 421 negative
Skin Sensitization 443
149 positive / 294 negative
Acute Oral Toxicity 261
86 positive / 175 negative
Acute Dermal Irritation 253
124 positive / 129 negative
Acute Dermal Toxicity 266
66 positive / 200 negative
Acute Aquatic Toxicity 268
50 positive / 218 negative
Acute Inhalation Toxicity 271
67 positive / 204 negative
Chronic Aquatic Toxicity 271
60 positive / 211 negative
Eye Irritation 259
133 positive / 126 negative
Mutagenicity 265
33 positive / 232 negative
Skin Sensitization 266
87 positive / 179 negative
Acute Oral Toxicity 323
175 positive / 148 negative
Acute Dermal Irritation 326
132 positive / 194 negative
Acute Dermal Toxicity 325
140 positive / 185 negative
Acute Aquatic Toxicity 320
188 positive / 132 negative
Acute Inhalation Toxicity 325
154 positive / 171 negative
Chronic Aquatic Toxicity 320
210 positive / 110 negative
Eye Irritation 325
123 positive / 202 negative
Mutagenicity 326
82 positive / 244 negative
Skin Sensitization 322
163 positive / 159 negative
Acute Oral Toxicity 101
59 positive / 42 negative
Acute Dermal Irritation 102
60 positive / 42 negative
Acute Dermal Toxicity 102
47 positive / 55 negative
Acute Aquatic Toxicity 101
32 positive / 69 negative
Acute Inhalation Toxicity 103
50 positive / 53 negative
Chronic Aquatic Toxicity 102
44 positive / 58 negative
Eye Irritation 100
60 positive / 40 negative
Mutagenicity 104
33 positive / 71 negative
Skin Sensitization 103
56 positive / 47 negative
Acute Oral Toxicity 67
25 positive / 42 negative
Acute Dermal Irritation 62
41 positive / 21 negative
Acute Dermal Toxicity 72
24 positive / 48 negative
Acute Aquatic Toxicity 70
8 positive / 62 negative
Acute Inhalation Toxicity 73
25 positive / 48 negative
Chronic Aquatic Toxicity 70
11 positive / 59 negative
Eye Irritation 63
46 positive / 17 negative
Mutagenicity 71
13 positive / 58 negative
Skin Sensitization 69
29 positive / 40 negative
Acute Oral Toxicity
1112 results
446 positive
666 negative
Acute Dermal Irritation
1071 results
547 positive
524 negative
Acute Dermal Toxicity
1146 results
350 positive
796 negative
Acute Aquatic Toxicity
1143 results
339 positive
804 negative
Acute Inhalation Toxicity
1171 results
389 positive
782 negative
Chronic Aquatic Toxicity
1143 results
413 positive
730 negative
Eye Irritation
1085
570 positive
515 negative
Mutagenicity
1161 results
204 positive
957 negative
Skin Sensitization
1124 results
460 positive
664 negative

Ultimately, a chemical list contains 1,215 chemicals had been generated. There are 10,935 Read-Across toxicological results correspond to the 1,215 chemicals. 1,215 chemicals had been divided into 8 manual curation categories. The 8 manual curation categories then condensed to 5 manual curation categories. After that, each category divided into 9 subcategories based on adverse health and environmental endpoints. In each subgroup, positive and negative results had been counted, and the quantities showed in above figure.

2.3 Validation

The validation process would be accomplished by comparing the Read-Across results with other toxicological results from other resources such as ECHA, OSHA, and other literature. It is important to know whether the Read-Across results are accurate, so the toxicological information for selected substances from other resources need to obtain and compared to show the consistency.

In order to validate the findings, 18 chemicals had been selected from the chemical list. The choice of chemicals was based on the following steps: Initially, using the sort and filter function in Microsoft Excel to separate chemicals based on their manual curation categories. Then, after separating, changing the order of the chemicals by descending prediction values. Highest 5 prediction values with very strong confidence levels have been selected from each manual curation category. Totally, 25 results had been picked. However, multiple results with different health endpoints could come from one chemical. For example, both eye irritation and acute dermal irritation for furfural are parts of the highest 5 prediction values in indirect food additive categories. Hence, corresponding to those 25 results, 18 chemicals had been selected. After deciding on the 18 chemicals, which would be used for validation purposes, all of their corresponding health endpoints, totally 162 results, have been recorded. A comparison table was made to compare the toxic evaluation results for Read-Across and other sources, including European Chemicals Agency, U. S. Environmental Protection Agency, Occupational Safety and Health Administration, European Food Safety Authority, and other literature. The table listed 18 chemicals with 162 toxicological assessment results. Whether the chemical will lead to the adverse health endpoints had been compared. For example, allyl cyclohexanepropionate has been identified as positive for acute aquatic toxicity in Read-Across, which means allyl cyclohexanepropionate will lead to acute aquatic toxicity based on Read-Across results. Meanwhile, other resources also point out allyl cyclohexanepropionate is very toxic to the

aquatic life and will lead to acute hazard. Hence, both results are match after comparison. Color codes had been assigned to each result. Green represented match, which means the results from other sources are match with the results from Read-Across, while red represented no-match, which means the results from other sources differ from the results from Read-Across. Colors also ranged from dark to light, which shows the confidence level provided by Read-Across.

3. Results

The purpose of the following comprehensive statistical analysis is testing the performance of Read-Across when predicting whether the chemical is a toxicants, and identifying which manual curation sub-group has the highest positive rate for each health and environmental endpoint; for each endpoint, the goal is testing the performance of Read-Across when suggesting positive results (having a threshold equal to or above 0.5) and causing adverse health endpoints, and for negative results (having a threshold below 0.5) and not cause adverse health endpoint. The aim is to show how accurate Read-Across is in food additives evaluation.

After statistical analysis, main findings were listed below.

1. It was rare to have two manual curation categories for one chemical. Generally, one chemical corresponds to one manual curation category.
2. There were more direct food additives and indirect food additives in the training data.
3. In this chemical list, there were more non-toxicants than toxicants, which was expected for food-related substances.
4. More results were founded at very strong and strong confidence level.
5. 83% of the Read-Across results selected for validation process match with the toxicological assessment results from other sources and literature.

3.1 Manual Curation Categorization

The purpose of this parts is identifying, which manual curation category has which percentage of chemicals with positive results. 1,215 chemicals had been copied in one separate sheet in Microsoft Excel, and each chemical had been assigned a code number based on their manual curation categorizations. Code number 1 represented direct food additive. Code number 2 represented indirect food additive. Code number 3 represented pesticide/residue. Code number 4 represented non-food. Code number 5 represented chemicals, which were not included in the manual curation. Code number 6 represented direct food additive + indirect food additive. Code number 7 represented direct food additive + pesticide/residue. Code number 8 represented indirect food additive + pesticide/residue. The SUMPRODUCT function in Excel has been used to count the number of appearances of each code number.

The counting results show that in the 1,215 chemicals, 455 chemicals are direct food additives. 19 chemicals are both direct food additives and indirect food additives. 28 chemicals are both direct food additives and pesticides/residues. 230 chemicals are indirect food additives. 35 chemicals are both indirect food additives and pesticides/residues. 106 chemicals are non-food. 266 chemicals are pesticides/residues. 76 chemicals are not included in the manual curation categorizations. The following table displayed each manual curation categorization in descending order. It is rare to have chemicals with two manual curations. In order to make the data easier to interpret, the counting results of indirect food additive + pesticide/residue, direct food additive + pesticide/residue, direct food additive + indirect food additive have been added to both the applying categories direct food additive, indirect food additive, pesticide/residue, respectively. After adding, the total counts and percentage exceeded 1,215 and 100% due to two manual curation categories had been counted twice.

Table 3.1 8 Manual Curation Categorization in 1,215 chemicals

Manual Curation Categorization	**Code number**	**Counts**	**Percentage**
Direct food additive	1	502	41%
Pesticide/residue	3	329	27%
Indirect food additive	2	284	23%
Non-food	4	106	9%
Not included in manual curation	5	76	6.26%
		1297	106%

Corresponding to the 1,215 chemicals, there are 10,935 Read-Across results with prediction value and confidence level. Filter and sort functions in Microsoft Excel have been used to count results in each manual curation category. In these 10,935 results, 4,095 results are for direct food additives. 171 results are both direct food additives and indirect food additives. 252 results are both direct food additives and pesticides/residues. 2,070 results are indirect food additives. 315 results are both indirect food additives and pesticides/residues. 954 results are non-food. 2,394 results are pesticides/residues. 684 results are not included in the manual curation categories. The following table displays each manual curation category in descending order. Because chemicals with two manual curations are still not a large portion, the counting results of indirect food additive + pesticide/residue, direct food additive + pesticide/residue, direct food additive +

indirect food additive have been added to both respective categories of direct food additive, indirect food additive, pesticide/residue either.

Table 3.2 8 Manual Curation Categorization in 10,935 results

Manual curation	**Counts**	**Percentage**
Direct food additive	4,518	41%
Pesticide/residue	2,961	27%
Indirect food additive	2,556	23%
Non-food	954	9%
Not included	684	6%
	10,935	106%

Even though Table 1 and Table 2 use different units, the datasets were the same. In consequence, the trend and percentages were basically the same. More than half of those 1,215 chemicals are the add-ups of direct food additives, indirect food additives, and pesticides/residues. Other manual curation categories only occupy a minor portion especially for direct food additive + indirect food additive, direct food additive + pesticide/residue, and indirect food additive + pesticide/residue. Based on this 1,215 chemicals list, chemicals that have only one manual curation category are far more common. It is actually rare that a chemical has two manual curation categories. For example, chemicals that can not only be a direct food additive but also be an indirect food additive represent only 2% of this chemical list. In that case, chemicals with two manual curation categories were not be showed in above tables and following tables and graphs.

3.2 In training data

In this part, the purpose is identifying, which manual curation category contains how many training data. “In training data” means these results were part of the training dataset to build the prediction model. These are also the data to be used for cross-validation when it becomes available as they have classifications based on animal studies. For example, $COC_1=NC=CN=C_1CC(C)C$ labeled as in training for eye irritation. This result indicates that the ability of Read-Across to predict whether $COC_1=NC=CN=C_1CC(C)C$ could trigger eye irritation was known and used to train the model; currently, the actual classification has not yet been disclosed by UL as they are proprietary; their release for the purpose of validation is under

negotiation. The SUMPRODUCT function in Microsoft Excel has been used to count the appearance of in training data for each health endpoint.

Table 3.3 In training data for 8 manual curation categories divided by 9 health endpoints

Positive + Negative in traning results for 8 manual curation divided by 9 health endpoints									
In Training Data	Direct food additive	Direct food additive, indirect food additive	Direct food additive, pesticide/residue	Indirect food additive	Indirect food additive, pesticide/residue	Non-food	Pesticide / residue	Not included in manual curation	
Acute oral toxicity	62	0	0	21	2	5	4	9	103
Acute dermal irritation	93	1	0	29	1	4	2	14	144
Acute dermal toxicity	40	0	0	17	1	4	3	4	69
Acute aquatic toxicity	38	0	0	14	2	5	7	6	72
Acute inhalation toxicity	22	0	0	12	1	3	3	3	44
Chronic aquatic toxicity	42	0	0	11	2	4	7	6	72
Eye irritation	81	1	0	23	1	6	3	13	128
Mutagenicity	26	0	0	18	1	2	2	5	54
Skin sensitization	58	1	0	15	2	3	5	7	91
	462	3	0	160	13	36	36	67	

Table 3.4 In training data for 5 manual curation categories

Manual curation	**In training results**	**Total results**	**Percentage**
Direct food additive	465	4,518	10%
Indirect food additive	176	2,556	7%
Non-food	36	954	4%
Pesticide/residue	49	2,961	2%
Not included	67	684	10%

Except for chemicals, which are not included in the manual curation categorization, direct food additives and indirect food additives contain the most in training results compared to other manual curation categories. Chemicals, which have two manual curation categories, typically have less in training results. This might be explained that these are the more broadly used substances with more likely having test data available.

Even though prediction values for health endpoint are provided by Read-Across, it would be better to exclude those results from further analysis. The following tables and comparisons only contain results, which have confidence levels of very strong & strong, moderate, weak, and exclude. In training results will not be included in the following calculations as the classification is at this stage not available to us as proprietary data. If they become available in the near future, they will be used for cross-validation and also included in the prevalence calculations for health and environmental endpoints.

3.3 Health and Environmental Endpoints

In this part, the purpose is identifying, which manual curation category has which positive rate for each health and environmental endpoint, and further characterize the performance of the Read-Across technology. In order to know this, the first step is counting the number of positive and negative results for each manual curation category and health or environmental endpoint. The total number for each health endpoint has been used as the denominator when calculating the percentage. For example, for acute oral toxicity, there is 11% positive direct food additives, and 11% was calculated by 125 divided by 1,112. The total percentage of direct food additive for acute oral toxicity is 40%, which is calculated by 440 divided by 1,112. For eye irritation, there is 12% positive indirect food additives, and the 12% was calculated by 133 divided by 1,085.

Table 3.5 Positive and negative results for each manual curation category and endpoint

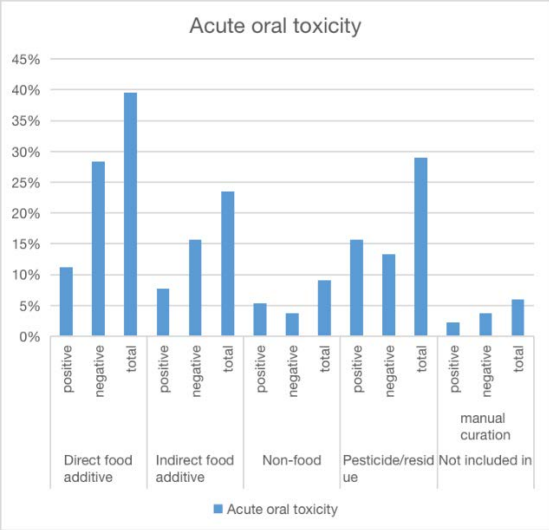
	Direct food additive			Indirect food additive			Non-food			Pesticide/residue			Not included in manual curation			ALL		
	positive	negative	total	positive	negative	total	positive	negative	total	positive	negative	total	positive	negative	total	positive	negative	total
Acute oral toxicity	125	315	440	86	175	261	59	42	101	175	148	323	25	42	67	446	666	1112
Acute dermal irritation	233	175	408	124	129	253	60	42	102	132	194	326	41	21	62	547	524	1071
Acute dermal toxicity	93	369	462	66	200	266	47	55	102	140	185	325	24	48	72	350	796	1146
Acute aquatic toxicity	76	388	464	50	218	268	32	69	101	188	132	320	8	62	70	339	804	1143
Acute inhalation toxicity	113	367	480	67	204	271	50	53	103	154	171	325	25	48	73	389	782	1171
Chronic aquatic toxicity	104	356	460	60	211	271	44	58	102	210	110	320	11	59	70	413	730	1143
Eye irritation	253	165	418	133	126	259	60	40	100	123	202	325	46	17	63	570	515	1085
Mutagenicity	55	421	476	33	232	265	33	71	104	82	244	326	13	58	71	204	957	1161
Skin sensitization	149	294	443	87	179	266	56	47	103	163	159	322	29	40	69	460	664	1124

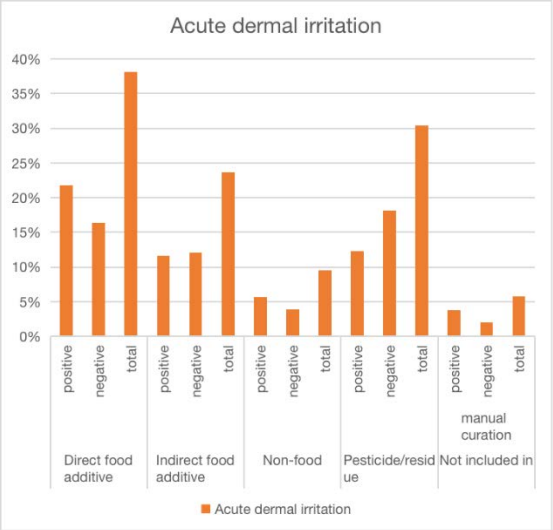
Table 3.6 Positive and negative results for each manual curation and endpoint in percentage

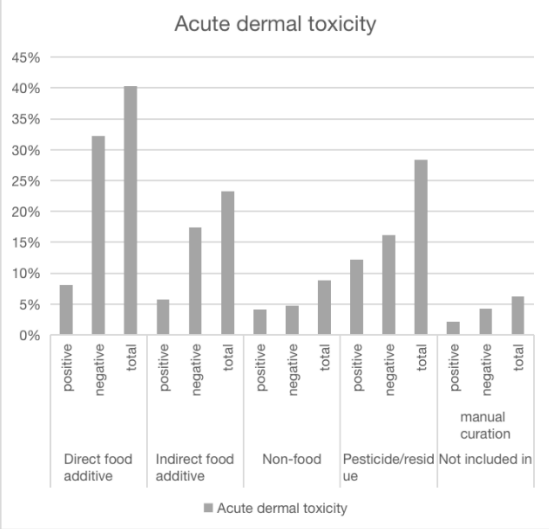
	Direct food additive			Indirect food additive			Non-food			Pesticide/residue			Not included in manual curation			ALL
	positive	negative	total	positive	negative	total	positive	negative	total	positive	negative	total	positive	negative	total	total
Acute oral toxicity	11%	28%	40%	8%	16%	23%	5%	4%	9%	16%	13%	29%	2%	4%	6%	1112
Acute dermal irritation	22%	16%	38%	12%	12%	24%	6%	4%	10%	12%	18%	30%	4%	2%	6%	1071
Acute dermal toxicity	8%	32%	40%	6%	17%	23%	4%	5%	9%	12%	16%	28%	2%	4%	6%	1146
Acute aquatic toxicity	7%	34%	41%	4%	19%	23%	3%	6%	9%	16%	12%	28%	1%	5%	6%	1143
Acute inhalation toxicity	10%	31%	41%	6%	17%	23%	4%	5%	9%	13%	15%	28%	2%	4%	6%	1171
Chronic aquatic toxicity	9%	31%	40%	5%	18%	24%	4%	5%	9%	18%	10%	28%	1%	5%	6%	1143
Eye irritation	23%	15%	39%	12%	12%	24%	6%	4%	9%	11%	19%	30%	4%	2%	6%	1085
Mutagenicity	5%	36%	41%	3%	20%	23%	3%	6%	9%	7%	21%	28%	1%	5%	6%	1161
Skin sensitization	13%	26%	39%	8%	16%	24%	5%	4%	9%	15%	14%	29%	3%	4%	6%	1124

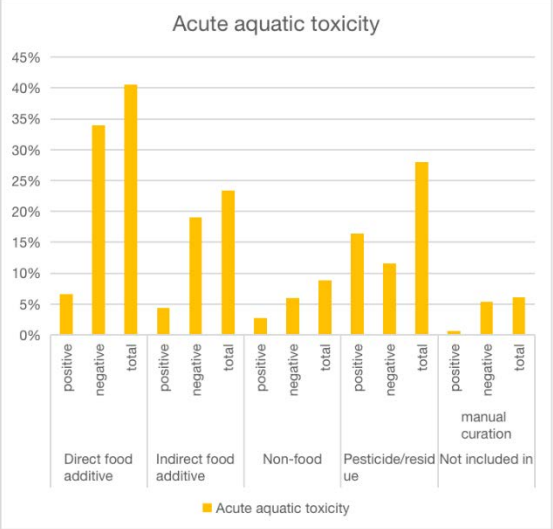
Based on these tables, several bar graphs have been generated. From the following bar graphs, it is easier to see the trends and to compare the results for each manual curation category. For example, in the acute oral toxicity bar graph, direct food additives, indirect food additives, and pesticides/residues have more results, and generally, negative results are more than positive results in direct food additive and indirect food additive for acute oral toxicity, which means non-toxicant results are more than toxicant results in this chemical list under the analysis of Read-Across. Not only for acute oral toxicity, acute dermal toxicity, acute aquatic toxicity, acute inhalation toxicity, chronic aquatic toxicity, mutagenicity and skin sensitization also have the similar distribution, which is negative results are more than positive results.

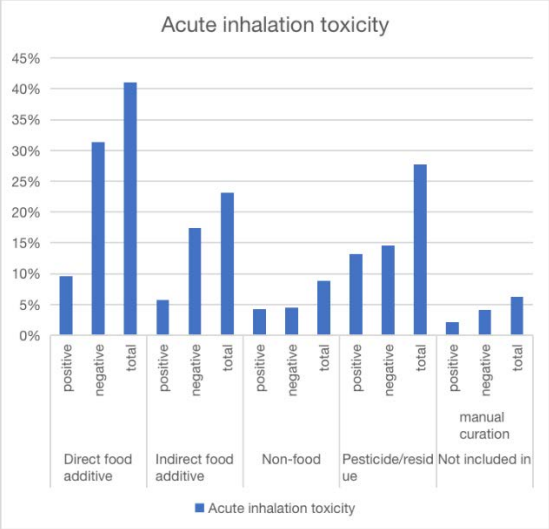
Figure 3.1 Positive and negative results for each manual curation category and endpoint

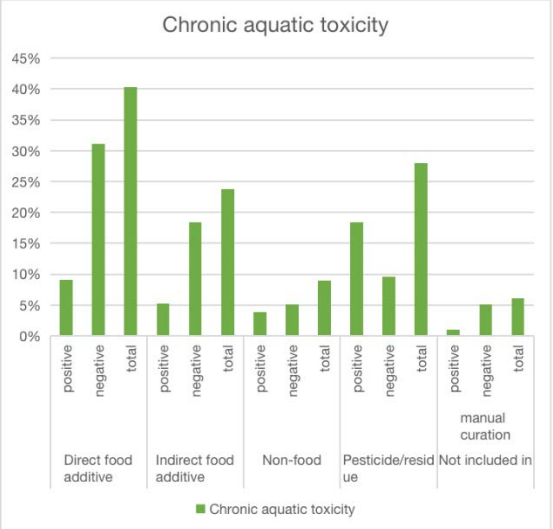

Acute oral toxicity
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute oral toxicity

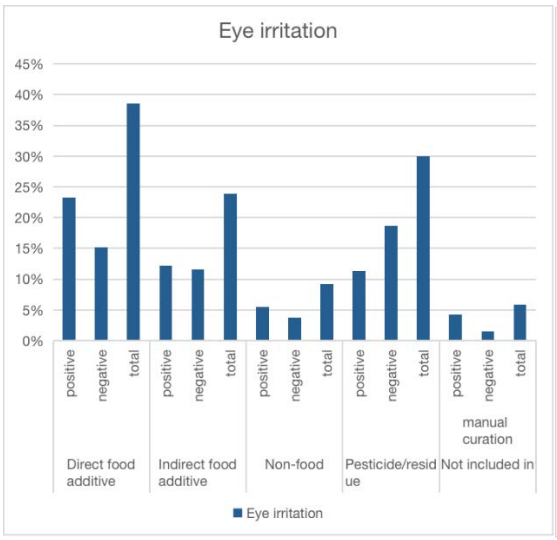

Acute dermal irritation
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute dermal irritation

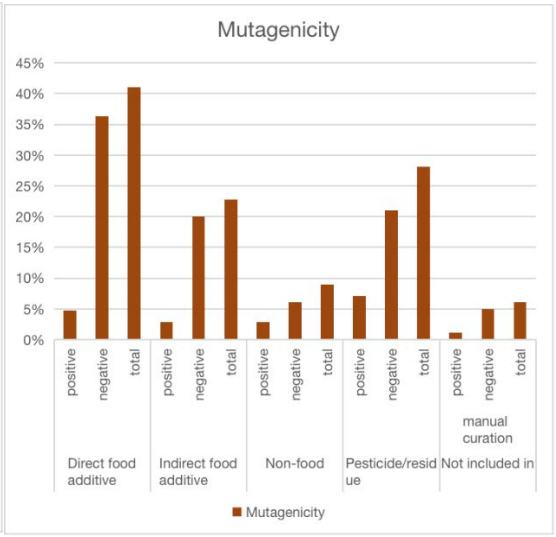

Acute dermal toxicity
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute dermal toxicity

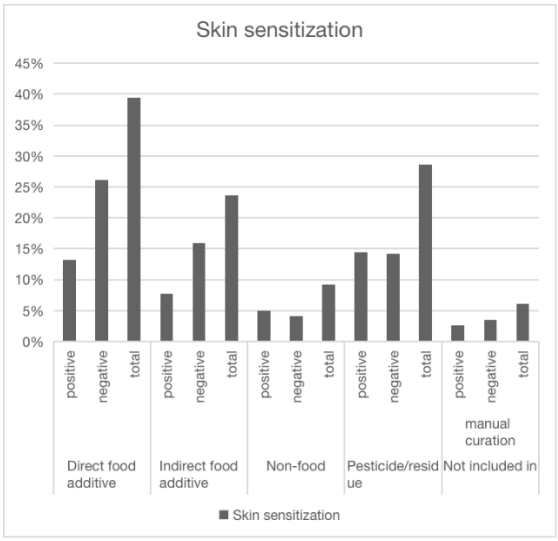

Acute aquatic toxicity
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute aquatic toxicity


Acute inhalation toxicity
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute inhalation toxicity


Chronic aquatic toxicity
45%
40%
35%
30%
25%
20%
15%
10%
5%
0%
positive
negative
total
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Chronic aquatic toxicity







The bar graph shows that for acute oral toxicity, acute dermal toxicity, acute aquatic toxicity, acute inhalation toxicity, chronic aquatic toxicity, mutagenicity, and skin sensitization, pesticide/residue have more positive results. For acute dermal irritation and eye irritation, direct food additives have more positive results. Generally, positive results were found for direct food additives, pesticides/residue, and indirect food additives throughout the manual curation categories, which corresponds to our previous data.

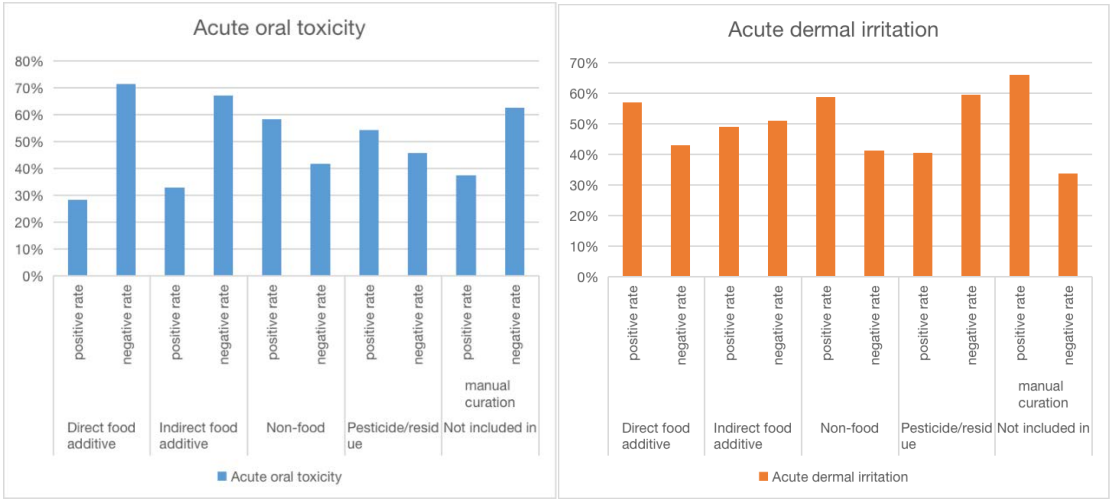
Table 3.7 Positive rate and negative rate

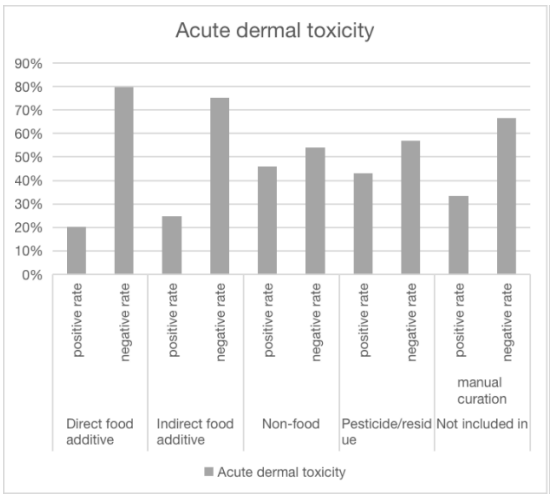
Positivity rate & Negativity rate										
	Direct food additive		Indirect food additive		Non-food		Pesticide/residue		Not included in manual curation	
	positive rate	negative rate	positive rate	negative rate	positive rate	negative rate	positive rate	negative rate	positive rate	negative rate
Acute oral toxicity	28%	72%	33%	67%	58%	42%	54%	46%	37%	63%
Acute dermal irritation	57%	43%	49%	51%	59%	41%	40%	60%	66%	34%
Acute dermal toxicity	20%	80%	25%	75%	46%	54%	43%	57%	33%	67%
Acute aquatic toxicity	16%	84%	19%	81%	32%	68%	59%	41%	11%	89%
Acute inhalation toxicity	24%	76%	25%	75%	49%	51%	47%	53%	34%	66%
Chronic aquatic toxicity	23%	77%	22%	78%	43%	57%	66%	34%	16%	84%
Eye irritation	61%	39%	51%	49%	60%	40%	38%	62%	73%	27%
Mutagenicity	12%	88%	12%	88%	32%	68%	25%	75%	18%	82%
Skin sensitization	34%	66%	33%	67%	54%	46%	51%	49%	42%	58%

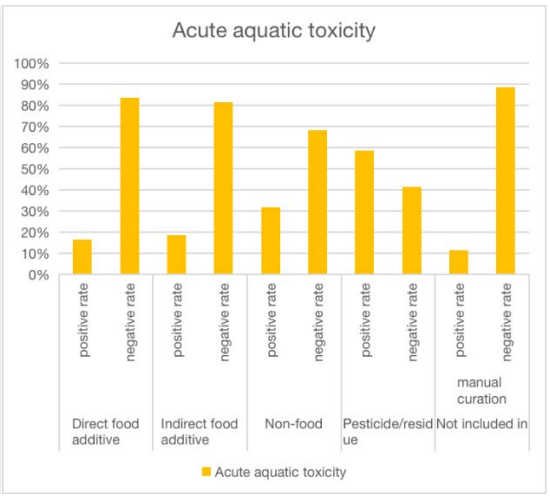
Based on Table 5, positive and negative rate could also be calculated. For example, in acute oral toxicity, the positive rate for direct food additives is 28%, which is calculated by 11% divided by 40%. 11% and 40% are retrieved from Table 5.

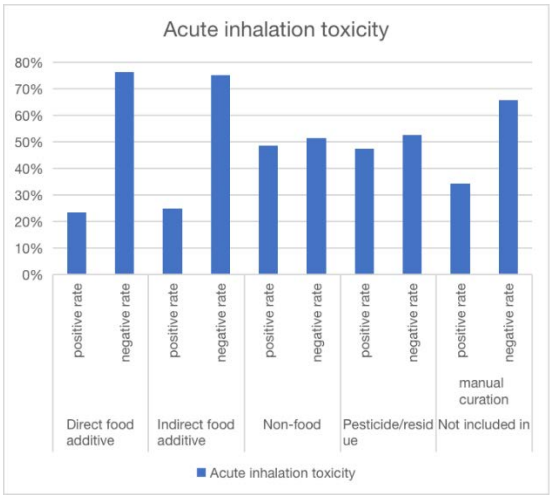
According to Table 6, bar charts could also be generated to compare positive and negative rate directly. For acute dermal irritation, it might be hard to determine whether there are more positive results than negative results. However, for other health endpoints, negative rates are generally greater than positive rates, which is corresponding to the findings in Table 4 and Table 5 and should be expected for food-related substances.

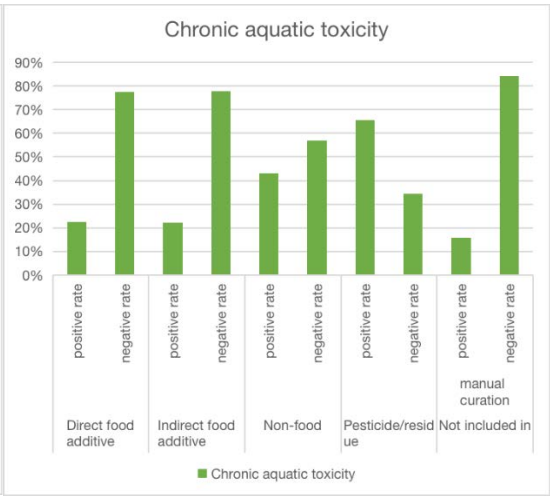
Figure 3.2 Positive rate and negative rate for each manual curation category and endpoint

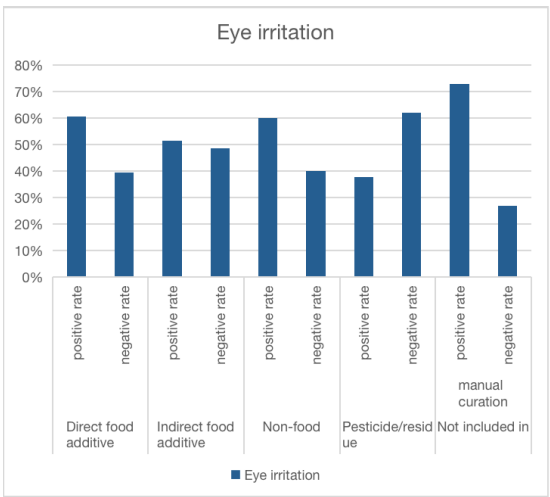


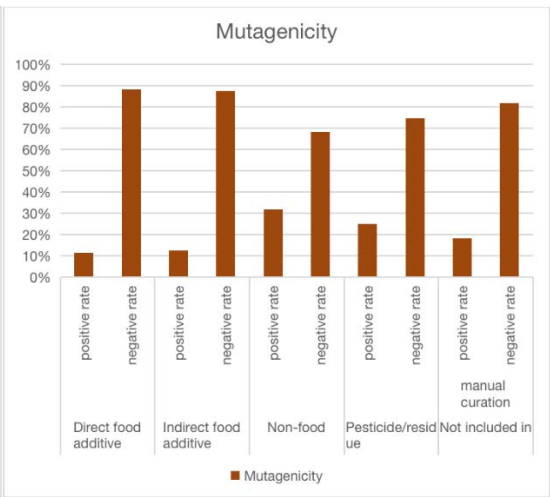

Acute dermal toxicity
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute dermal toxicity

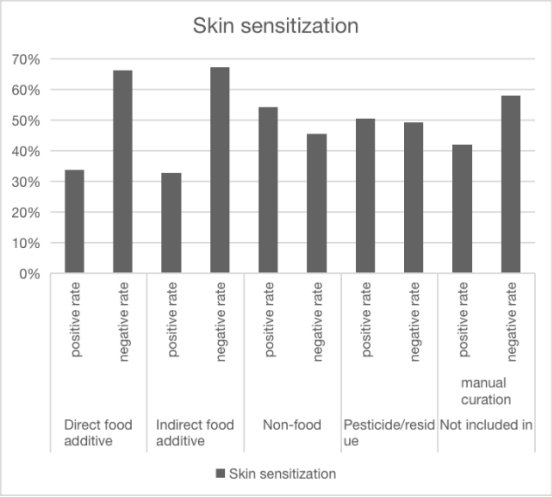

Acute aquatic toxicity
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute aquatic toxicity


Acute inhalation toxicity
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Acute inhalation toxicity


Chronic aquatic toxicity
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Chronic aquatic toxicity


Eye irritation
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Eye irritation


Mutagenicity
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
positive rate
negative rate
manual curation
Direct food additive
Indirect food additive
Non-food
Pesticide/residue
Not included in
Mutagenicity

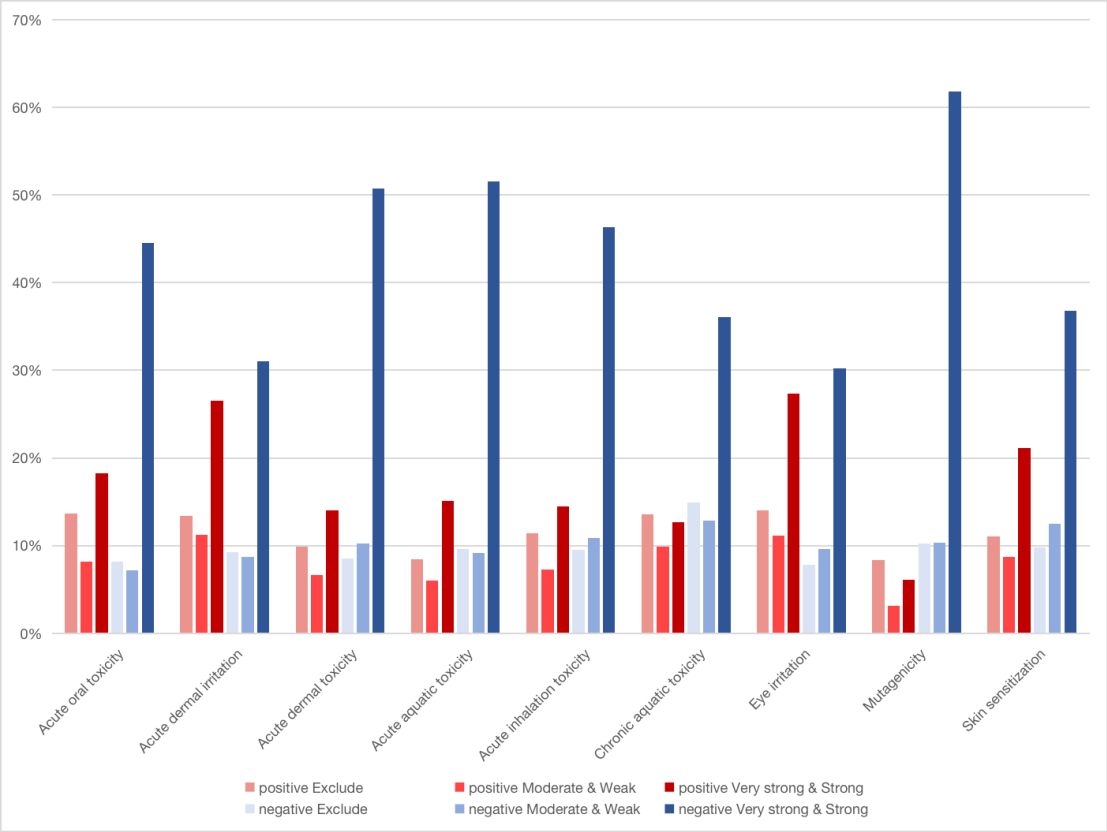


3.4 Confidence Level

In this part, the purpose is identifying how confident the program is when targeting different health and environmental endpoints. There are three confidence levels, the first one is very strong & strong, which means the program is confident about the prediction value. The second one is moderate & weak, which means the program has low confidence about the prediction value. The last one is “exclude”, which means the structure does not fit ULCT rules and thus we cannot obtain the confidence level.

Divided by health endpoints, this part shows the accuracy of the predicting results. The following graph provides the total positive and negative results at the different confidence levels for the specific health and environmental endpoint. By comparing the positive and negative numbers, it is able to know whether the program is good at predicting the positive results or negative results, and also able to know how confident the program is when treating the specific type of endpoint.

Figure 3.3 Positive and Negative results for 3 confidence levels



Red bars mean the results are positive, which is a toxicant, while blue bars mean the results are negative, which is a non-toxicant. Darker color means Read-Across is more confident about the prediction results, while lighter color means Read-Across is less confident about the prediction results. From the bar graph, it is simple to see the chemical is non-toxicant with high confidence. It is also capable of identifying chemicals that could lead to adverse health endpoints with high confidence, but, in general, Read-Across will have better performance targeting non-toxicants than toxicants in this 1,215-chemical list. In addition, there are much more results in very strong and strong positive and very strong and strong negative comparing to moderate and weak strong and moderate and weak negative, which indicates Read-Across could be a promising tool when predicting the adverse health and environmental effects for chemicals in this list.

3.5 Validation

In this part, by comparing the toxicological assessment results from Read-Across and other sources, the validation process could be completed. It is important to know whether Read-Across could perform in an accurate way, so comparing the results obtained from Read-Across to toxicological assessment results from other sources, such as ECHA, OSHA and other literature

could be a useful and feasible method. It is able to show Read-Across is accurate enough compare to traditional testing methods.

The highest 5 prediction values with very strong confidence level in 5 manual curation categories (direct food additives, indirect food additives, pesticides/residues, non-food, not included in manual curation categorization) had been selected. 18 chemicals are corresponding to these 25 prediction values. Those 18 chemicals included allyl cyclohexanepropionate, α-Phellandrene, methyl butyrate, 3-(Methylthio)propyl isothiocyanate, pentachloropyridine, furfural, 2,4-diaminotoluene, dichlorobenzene, coumaphos, coumatetralyl, sulfotep, 2.4-D-1-butyl ester, terbufos, tefluthrin, deltamethrin, cypermethrin, fenvalerate, and 2,5-dimethulfuran. 18 chemicals with 162 results for 9 health and environmental endpoints had been selected for the validation process.

Some health and environmental endpoints for some chemicals cannot found available data from other sources and literature. For example, α-Phellandrene has a pretty high prediction value, around 0.994, for acute aquatic toxicity in Read-Across, and Read-Across shows high confidence level for this prediction value. However, in other sources or literature, no available data has been provided for whether α-Phellandrene would lead to acute aquatic toxicity and be harmful to aquatic life. This kind of results had been excluded from the validation process, and results marked as “exclude” in Read-Across do not use for validation, neither. Finally, 123 results are eligible for validation.

Table 3.8 Comparison table

Chemical	CASRNs	Manual curation categorization	Health and environmental endpoints	Read-Across results	Other sources	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Acute aquatic toxicity	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Acute oral toxicity	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Chronic aquatic toxicity	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Skin sensitization	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Acute inhalation toxicity	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Acute dermal toxicity	√	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Acute dermal irritation	/	√	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Eye irritation	x	/	
allyl cyclohexanepropionate	2705-87-5	Direct food additive	Mutagenicity	x	x	
α-phellandrene	99-83-2	Direct food additive	Acute aquatic toxicity	√	/	
α-phellandrene	99-83-2	Direct food additive	Skin sensitization	√	√	
α-phellandrene	99-83-2	Direct food additive	Chronic aquatic toxicity	√	√	
α-phellandrene	99-83-2	Direct food additive	Acute dermal irritation	x	√	
α-phellandrene	99-83-2	Direct food additive	Acute inhalation toxicity	x	√	
α-phellandrene	99-83-2	Direct food additive	Mutagenicity	x	/	

α-phellandrene	99-83-2	Direct food additive	Eye irritation	X	√	
α-phellandrene	99-83-2	Direct food additive	Acute dermal toxicity	X	/	
α-phellandrene	99-83-2	Direct food additive	Acute oral toxicity	X	√	
Methyl butyrate	623-42-7	Direct food additive	Eye irritation	√	√	
Methyl butyrate	623-42-7	Direct food additive	Acute dermal irritation	√	√	
Methyl butyrate	623-42-7	Direct food additive	Acute inhalation toxicity	X	X	
Methyl butyrate	623-42-7	Direct food additive	Mutagenicity	X	/	
Methyl butyrate	623-42-7	Direct food additive	Acute dermal toxicity	X	/	
Methyl butyrate	623-42-7	Direct food additive	Skin sensitization	X	X	
Methyl butyrate	623-42-7	Direct food additive	Chronic aquatic toxicity	X	/	
Methyl butyrate	623-42-7	Direct food additive	Acute aquatic toxicity	X	/	
Methyl butyrate	623-42-7	Direct food additive	Acute oral toxicity	X	X	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Acute aquatic toxicity	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Acute dermal irritation	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Acute oral toxicity	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Eye irritation	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Chronic aquatic toxicity	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Skin sensitization	√	/	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Acute inhalation toxicity	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Acute dermal toxicity	√	√	
3-(Methylthio)propyl isothiocyanate	505-79-3	Direct food additive	Mutagenicity	X	/	
Pentachloropyridine	2176-62-7	Indirect food additive	Acute aquatic toxicity	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Acute inhalation toxicity	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Acute dermal irritation	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Acute dermal toxicity	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Eye irritation	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Skin sensitization	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Chronic aquatic toxicity	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Acute oral toxicity	√	√	
Pentachloropyridine	2176-62-7	Indirect food additive	Mutagenicity	X	X	
Furfural	98-01-1	Indirect food additive	Eye irritation	√	√	
Furfural	98-01-1	Indirect food additive	Acute dermal irritation	√	√	
Furfural	98-01-1	Indirect food additive	Skin sensitization	√	√	
Furfural	98-01-1	Indirect food additive	Acute inhalation toxicity	√	√	
Furfural	98-01-1	Indirect food additive	Acute dermal toxicity	√	√	
Furfural	98-01-1	Indirect food additive	Acute oral toxicity	√	√	
Furfural	98-01-1	Indirect food additive	Mutagenicity	√	√	
Furfural	98-01-1	Indirect food additive	Acute aquatic toxicity	/	√	
Furfural	98-01-1	Indirect food additive	Chronic aquatic toxicity	/	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Acute dermal toxicity	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Mutagenicity	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Skin sensitization	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Acute inhalation toxicity	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Acute oral toxicity	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Acute dermal irritation	√	√	

2,4-Diaminotoluene	95-80-7	Indirect food additive	Acute aquatic toxicity	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Eye irritation	√	√	
2,4-Diaminotoluene	95-80-7	Indirect food additive	Chronic aquatic toxicity	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Acute dermal irritation	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Eye irritation	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Acute dermal toxicity	√	X	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Mutagenicity	√	X	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Acute inhalation toxicity	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Acute oral toxicity	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Skin sensitization	√	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Acute aquatic toxicity	/	√	
Dichlorobenzene	106-46-7	Indirect food additive, pesticides/residues	Chronic aquatic toxicity	X	√	
Coumaphos	56-72-4	Non-food	Acute aquatic toxicity	√	√	
Coumaphos	56-72-4	Non-food	Acute inhalation toxicity	√	√	
Coumaphos	56-72-4	Non-food	Acute dermal toxicity	√	√	
Coumaphos	56-72-4	Non-food	Chronic aquatic toxicity	√	√	
Coumaphos	56-72-4	Non-food	Acute oral toxicity	√	√	
Coumaphos	56-72-4	Non-food	Skin sensitization	√	√	
Coumaphos	56-72-4	Non-food	Mutagenicity	/	X	
Coumaphos	56-72-4	Non-food	Acute dermal irritation	X	√	
Coumaphos	56-72-4	Non-food	Eye irritation	X	√	
Coumatetralyl	5836-29-3	Non-food	Acute aquatic toxicity	√	√	
Coumatetralyl	5836-29-3	Non-food	Chronic aquatic toxicity	√	√	
Coumatetralyl	5836-29-3	Non-food	Acute oral toxicity	√	√	
Coumatetralyl	5836-29-3	Non-food	Skin sensitization	√	X	
Coumatetralyl	5836-29-3	Non-food	Acute dermal irritation	√	√	
Coumatetralyl	5836-29-3	Non-food	Acute inhalation toxicity	√	√	
Coumatetralyl	5836-29-3	Non-food	Acute dermal toxicity	√	√	
Coumatetralyl	5836-29-3	Non-food	Eye irritation	√	√	
Coumatetralyl	5836-29-3	Non-food	Mutagenicity	/	√	
Sulfotep	3689-24-5	Non-food	Acute aquatic toxicity	√	√	
Sulfotep	3689-24-5	Non-food	Chronic aquatic toxicity	√	√	
Sulfotep	3689-24-5	Non-food	Acute oral toxicity	√	√	
Sulfotep	3689-24-5	Non-food	Acute dermal toxicity	√	√	
Sulfotep	3689-24-5	Non-food	Acute inhalation toxicity	√	√	
Sulfotep	3689-24-5	Non-food	Skin sensitization	/	√	
Sulfotep	3689-24-5	Non-food	Acute dermal irritation	/	√	
Sulfotep	3689-24-5	Non-food	Eye irritation	X	√	
Sulfotep	3689-24-5	Non-food	Mutagenicity	X	X	
2,4-D-1-butyl ester	94-80-4	Non-food	Acute aquatic toxicity	√	√	
2,4-D-1-butyl ester	94-80-4	Non-food	Acute inhalation toxicity	√	/	
2,4-D-1-butyl ester	94-80-4	Non-food	Skin sensitization	√	√	
2,4-D-1-butyl ester	94-80-4	Non-food	Acute oral toxicity	√	√	
2,4-D-1-butyl ester	94-80-4	Non-food	Acute dermal toxicity	√	/	
2,4-D-1-butyl ester	94-80-4	Non-food	Chronic aquatic toxicity	√	√	

2,4-D-1-butyl ester	94-80-4	Non-food	Acute dermal irritation	√	√	
2,4-D-1-butyl ester	94-80-4	Non-food	Eye irritation	/	√	
2,4-D-1-butyl ester	94-80-4	Non-food	Mutagenicity	X	X	
Terbufos	13071-79-9	Pesticides/residues	Acute aquatic toxicity	√	√	
Terbufos	13071-79-9	Pesticides/residues	Chronic aquatic toxicity	√	√	
Terbufos	13071-79-9	Pesticides/residues	Acute oral toxicity	/	√	
Terbufos	13071-79-9	Pesticides/residues	Acute inhalation toxicity	/	√	
Terbufos	13071-79-9	Pesticides/residues	Skin sensitization	/	/	
Terbufos	13071-79-9	Pesticides/residues	Acute dermal toxicity	/	√	
Terbufos	13071-79-9	Pesticides/residues	Acute dermal irritation	/	√	
Terbufos	13071-79-9	Pesticides/residues	Mutagenicity	X	X	
Terbufos	13071-79-9	Pesticides/residues	Eye irritation	X	√	
Tefluthrin	79538-32-2	Pesticides/residues	Acute aquatic toxicity	√	√	
Tefluthrin	79538-32-2	Pesticides/residues	Chronic aquatic toxicity	√	√	
Tefluthrin	79538-32-2	Pesticides/residues	Acute inhalation toxicity	√	√	
Tefluthrin	79538-32-2	Pesticides/residues	Acute dermal toxicity	√	√	
Tefluthrin	79538-32-2	Pesticides/residues	Acute oral toxicity	/	√	
Tefluthrin	79538-32-2	Pesticides/residues	Skin sensitization	X	X	
Tefluthrin	79538-32-2	Pesticides/residues	Mutagenicity	X	X	
Tefluthrin	79538-32-2	Pesticides/residues	Acute dermal irritation	X	√	
Tefluthrin	79538-32-2	Pesticides/residues	Eye irritation	X	√	
Deltamethrin	52918-63-5	Pesticides/residues	Acute aquatic toxicity	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Skin sensitization	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Chronic aquatic toxicity	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Acute oral toxicity	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Acute dermal irritation	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Acute inhalation toxicity	√	√	
Deltamethrin	52918-63-5	Pesticides/residues	Acute dermal toxicity	/	/	
Deltamethrin	52918-63-5	Pesticides/residues	Eye irritation	/	√	
Deltamethrin	52918-63-5	Pesticides/residues	Mutagenicity	/	/	
Cypermethrin	52315-07-8	Pesticides/residues	Acute aquatic toxicity	√	√	
Cypermethrin	52315-07-8	Pesticides/residues	Chronic aquatic toxicity	√	√	
Cypermethrin	52315-07-8	Pesticides/residues	Acute oral toxicity	/	√	
Cypermethrin	52315-07-8	Pesticides/residues	Acute inhalation toxicity	/	√	
Cypermethrin	52315-07-8	Pesticides/residues	Skin sensitization	X	√	
Cypermethrin	52315-07-8	Pesticides/residues	Acute dermal toxicity	X	/	
Cypermethrin	52315-07-8	Pesticides/residues	Acute dermal irritation	X	√	
Cypermethrin	52315-07-8	Pesticides/residues	Mutagenicity	X	X	
Cypermethrin	52315-07-8	Pesticides/residues	Eye irritation	X	√	
Fenvalerate	51630-58-1	Pesticides/residues	Acute aquatic toxicity	√	√	
Fenvalerate	51630-58-1	Pesticides/residues	Chronic aquatic toxicity	√	√	
Fenvalerate	51630-58-1	Pesticides/residues	Acute oral toxicity	/	√	
Fenvalerate	51630-58-1	Pesticides/residues	Skin sensitization	X	√	

Fenvalerate	51630-58-1	Pesticides/residues	Acute inhalation toxicity	x	/	
Fenvalerate	51630-58-1	Pesticides/residues	Acute dermal irritation	x	√	
Fenvalerate	51630-58-1	Pesticides/residues	Mutagenicity	x	x	
Fenvalerate	51630-58-1	Pesticides/residues	Acute dermal toxicity	x	/	
Fenvalerate	51630-58-1	Pesticides/residues	Eye irritation	x	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Acute dermal irritation	√	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Eye irritation	√	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Acute inhalation toxicity	√	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Acute dermal toxicity	√	/	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Skin sensitization	√	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Acute oral toxicity	√	√	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Mutagenicity	√	x	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Acute aquatic toxicity	/	/	
2.5-Dimethylfuran	625-86-5	Not included in manual curation	Chronic aquatic toxicity	/	/	

The above table is the simplified version of comparison table. Detailed version with descriptions for Read-Across results and results from other sources is in the supplementary material. The first four columns listed chemical names, CASRNs, manual curation categories, and health and environmental endpoints. For each chemical, health and environmental endpoints with higher prediction values are on the top, while health and environmental endpoints with lower prediction values are below. The fifth column listed the toxicological assessment results from Read-Across. Check (√) mark represents prediction values above the threshold 0.5, which is able to lead this kind of health and environmental endpoints. Cross (x) mark represents prediction values below the threshold 0.5, which fail to cause this kind of health and environmental endpoints. Slash (/) mark represents the result provided by Read-Across for this health and environmental endpoint is “exclude”. The sixth column listed the toxicological assessment results from other sources and literature. Check (√) mark represents other sources or literature do report this chemical will lead to this kind of health and environmental endpoint. Cross (x) mark represents other sources or literature report this chemical will not lead to this kind of health and environmental endpoint. Slash (/) mark represents toxicological assessment results failed to be found in other sources or literature, or the data is not available. The last column is the color scheme for every pair of result. Dark green represents the results from Read-Across and other sources are match, and Read-Across shows high confidence level for the prediction values. Light green represents the results from Read-Across and other sources are match, but Read-Across shows low confidence level for the prediction values. Dark red represents the result are not match, but Read-Across shows high confidence level for the prediction values. Light red

represents the result are not match, and Read-Across does shows low confidence level for the prediction values. Pink represents the results are not match. While Read-Across shows high confidence level for the prediction values and note the chemical will not lead to this type of health and environmental endpoint, other sources do report health and environmental endpoint could be found, but the extend is merely mild or slight. In addition, pink is a subset of dark red. Yellow represents the confidence level is marked as “exclude” in Read-Across. For results which has been marked as negative in Read-Across, if toxicological assessment results cannot be found in other sources or literature, it had been marked as blue. For results which has been marked as positive in Read-Across, if taxological assessment results cannot be found in other sources or literature, it had been marked as orange. Pairs of results which marked as yellow, blue and orange are those results which are not eligible for validation, and the quantities for yellow, blue and orange are 23, 11 and 5 respectively.

In conclusion, there are totally 162 pairs of results, but only 123 pairs of results are eligible for validation. Paris of results marked as dark green, light green, dark red, light red and pink are 86, 16, 10, 7 and 4 respectively. The percentage for match and not match are 83% and 17% calculated by 102 divided by 123 and 21 divided by 123.

4. Discussion

Artificial intelligence and big data have paved the way for the safety evaluation of food additives. As the increasing use of a variety of food additives, more and more chemicals need to be evaluated for safety issues before adding to food and entering the market. Considering the materials, money, and time which will be invested in traditional animal testing, Read-Across becomes a promising method for safety evaluation as it would be convenient and accurate enough. A total of 1,215 chemicals were used for descriptive statistical analysis, and these 1.215 chemicals are based on Dr. Karmaus' 2016 and 2017 works and Read-Across results provided by Dr. Luechtefeld. Chemicals were analyzed comprehensively to see the entire trends, but also divided into subgroups based on manual curation categories, health and environmental endpoints, and confidence levels to show the performance in each subgroup. Results showed that one chemical, generally, corresponding to one manual curation category, but chemicals could also be indirect food additive or pesticides/residues as they are direct food additives. In training results, which are proprietary, could be used for cross-validation once the data have been released. Chemicals with multiple manual curation categories typically have less in training results, which could be explained as these chemicals are broadly used substances in a variety of fields with more available data. This 1,215-chemical list contain more non-toxicants as Read-Across results showed more negative results when analyzing the chemicals based on health and environmental endpoints subgroup. The analysis of confidence level subgroups also showed Read-Across has very high confidence level towards those results, which have been labeled as non-toxicants. In addition, the analysis of confidence level subgroups also provides high confidence level for toxicants, which shows the good performance of Read-Across when conducting safety evaluation.

Direct food additives are considered as the highest exposure likelihood from food use followed with indirect food additive and pesticides/residues (Karmaus et al., 2017). This categorization could help explained why direct food additives, indirect food additives, and pesticides/residues represent the largest percentage in the 1,215-chemical list.

During the validation process, a comparison table was made for 18 chemicals with 162 results, but only 123 results are eligible for validation. 23 results labeled as "exclude" in Read-

Across. 16 results fail to find available data outside Read-Across, so these 39 results would be inappropriate to be included in validation process.

Table 4.1 Chemicals used for validation separated by color scheme

	Positive	Percentage*	Negative	Percentage*
Dark green	78	63%	8	7%
Light green	12	10%	4	3%
Dark red	3	2%	7	6%
Light red	1	1%	6	5%
Pink	0	0%	4	3%

*The denominator used for calculating percentage is 123, which is the eligible for validation counts for green & red

In the 123 results which is eligible for validation, the evaluation from other sources for 102 results were matched with Read-Across, while 21 results were no-match. The corresponding percentages for match and no-match are 83% and 17%, respectively. In the comparison table, matched results were labeled as dark green and light green color, and dark green represents Read-Across has a high confidence level for the prediction values while light green represents Read-Across has low confidence level for the prediction. Results, which are no-match, were labeled as dark red, light red, and pink in the comparison table. Dark red means evaluation and information from other sources differ from the Read-Across results though Read-Across has high confidence level for the prediction. Light red means evaluation and information from other sources differ from Read-Across results and Read-Across does have low confidence level for the prediction. Results labeled as pink also has very strong confidence level in Read-Across evaluation but labeled as slightly or mildly hazards in other sources. The results have also been separated by the 0.5 threshold and showed that more toxicants were be used in the validation process. Overall, it shows the good performance of Read-Across as the safety evaluation results are highly matched. It has to be noted, though, that substances with strong predicted effect and confidence were selected.

Here, what needs to be more attention paid to is that acute aquatic toxicity always has the highest prediction values for pesticides/residues manual curation category in the selected chemicals used for validation. Pesticides/residues enter the environmental and contaminate the water system. It is toxic to aquatic life and cause long-lasting effects. Except for chemicals categorized as pesticides/residues, some chemicals categorized as direct food additive, indirect food additive, and non-food also have highest prediction values for acute aquatic toxicity.

Karmaus categorized chemicals which do not use in the U.S. anymore or foreign-use pesticides, drugs, components of cosmetics, and industrial chemicals as non-food (Karmaus et al., 2016), which means non-food could contain pesticides too. The exposure route for non-food could be diverse due to the multiple usages, and lead to the contamination of water systems. However, for direct food additives and indirect food additives, it would be interesting to investigate the reason behind the highest prediction values for acute aquatic toxicity endpoint.

It is important to remember that chemicals mentioned in this research are mainly food-relevant chemicals currently used in the Unites States. Even though chemicals in non-food manual curation category do include food-relevant chemicals used in foreign countries, it might not be a representative list as the quantity is not large enough. In addition, there still are data gaps. For Read-Across, the detailed information for in training data, which could be used to conduct cross-validation, are not currently available. Several chemicals also do not have too much information and available data outside Read-Across, which prevent us from validating the prediction results for that health and environmental endpoints. For example, 3-(Methylthio)propyl isothiocyanate has about 0.9 prediction value in Read-Across for skin sensitization. 0.9 is a pretty high prediction value, which indicated 3-(Methylthio)propyl isothiocyanate has a high likelihood for inducing skin sensitization as a direct food additive, flavoring agent. However, during the process of searching information in other sources, there is not too much research conducted on this chemical to evaluate its skin sensitization probability, or data had been labeled as not available. Conflicting results also existed in data, which had high confidence level in Read-Across: α-Phellandrene is typical in the 18 chemicals used for validation. Only α-Phellandrene has more no-match results than matched results. 5 health and environmental endpoints have prediction value less than 0.15, which should be considered as a non-toxicant, labeled as fatal or harmful in other sources.

The 1,215 chemicals used in this analysis are came from Dr. Karmaus's work, and Dr. Karmaus used Dr. Dionisio's work. In Dr. Dionisio's work, "food" chemicals had been extracted based on specific word "food". However, some substances may not include "food" in the description field. It is likely that some important substances had been excluded from the analyses due to this filtering method.

5. Conclusions

This study demonstrates that Read-Across could be a useful and trustworthy tool for food safety evaluation as the use of for food additives becomes more frequent and diverse. This is a significant step in benefitting public health from consuming ingredients, which could have potential adverse health and environmental effects and achieve the goal of 3Rs (replacement, reduction, and refinement) of animal use in toxicological testing. Future studies should focus on chemicals, which have not been subjected to full studies for those 9 health and environmental endpoints (acute oral toxicity, acute dermal irritation, acute dermal toxicity, chronic aquatic toxicity, eye irritation, acute inhalation toxicity, acute aquatic toxicity, mutagenicity, skin sensitization) to gain more available data and then broaden the chemical list eligible for validation. Cross-validation should also be done once the “in training” data become available.

References

1. Beginning in October 1999, FDA began to authorize food contact substances including indirect food additives through the food contact notification process. This notification process does not result in a regulation in 21 CFR. FDA does list information on "effective" (i.e., those that complete the authorization process) notifications for food contact substances on its internet site. http://www.fda.gov/Food/IngredientsPackagingLabeling/PackagingFCS/Notifications/ucm116567.htm. Additionally, FDA has operated an exemption process for listing indirect food additives in 21 CFR since 1995. See Code of Federal Regulations, Title 21, Section 170.39 (Threshold of regulation for substances used in food-contact articles) (2013).
2. Code of Federal Regulations, Title 21, Part 58: Good Laboratory Practice for Nonclinical Laboratory Studies (2013).
3. Dionisio, K. L., Frame, A. M., Goldsmith, M.-R., Wambaugh, J. F., Liddell, A., Cathey, T., Smith, D., Vail, J., Ernstoff, A. S., Fantke, P., Jolliet, O., & Judson, R. S. (2015). Exploring consumer exposure pathways and patterns of use for chemicals in the environment. Toxicology Reports, 2, 228-237. http://doi.org/10.1016/j.toxrep.2014.12.009
4. FDA. (2000). Toxicological Principles for the Safety Assessment of Food Ingredients Redbook 2000.
5. FDA. (2017). Code of Federal Regulation (CFR) Citations for Color Additives, Food Ingredients and Packaging. https://www.fda.gov/food/ingredients-additives-gras-packaging-guidance-documents-regulatory-information/code-federal-regulations-cfr-citations-color-additives-food-ingredients-and-packaging
6. FDA. (2017). High-Intensity Sweeteners. https://www.fda.gov/food/food-additives-petitions/high-intensity-sweeteners
7. FJODOROVA, N., NOVICH, M., VRACHKO, M., SMIRNOV, V., KHARCHEVNIKOVA, N., ZHOLDAKOVA, Z., NOVIKOV, S., SKVORTSOVA, N., FILIMONOV, D., POROIKOV, V., & BENFENATI, E. (2008). Directions in QSAR Modeling for Regulatory Uses in OECD Member Countries, EU and in Russia. Journal of Environmental Science and Health, Part C, 26(2), 201–236. https://doi.org/10.1080/10590500802135578
8. GAO. (2010). Food Safety: FDA should strengthen its oversight of food ingredients determined to be generally recognized as safe (GRAS). http://www.gao.gov/products/GAO-10-246
9. Hartung, T. (2016). Making big sense from big data in toxicology by read-across. ALTEX, 83–93. https://doi.org/10.14573/altex.1603091
10. Hartung, T. (2017). Thresholds of Toxicological Concern – Setting a threshold for testing below which there is little concern. ALTEX, 331–351. https://doi.org/10.14573/altex.1707011
11. Hartung, T. (2018). Rebooting the generally recognized as safe (GRAS) approach for food additive safety in the US. ALTEX, 3–25. https://doi.org/10.14573/altex.1712181
12. Hartung, T. and Rovida, C. (2009). Food for thought … on in silico methods in toxicology. ALTEX, 155–166. https://doi.org/10.14573/altex.2009.3.155
13. Karmaus, A. L., Filer, D. L., Martin, M. T., & Houck, K. A. (2016). Evaluation of food-relevant chemicals in the ToxCast high-throughput screening program. Food and Chemical Toxicology, 92, 188–196. https://doi.org/10.1016/j.fct.2016.04.012
14. Karmaus, A. L., Trautman, T. D., Krishan, M., Filer, D. L., & Fix, L. A. (2017). Curation of

food-relevant chemicals in ToxCast. Food and Chemical Toxicology, 103, 174–182. https://doi.org/10.1016/j.fct.2017.03.006
15. Luechtefeld, T. and Hartung, T. (2017). Computational approaches to chemical hazard assessment. ALTEX, 459–478. https://doi.org/10.14573/altex.1710141
16. Luechtefeld, T., Marsh, D., Rowlands, C., & Hartung, T. (2018). Machine Learning of Toxicological Big Data Enables Read-Across Structure Activity Relationships (RASAR) Outperforming Animal Test Reproducibility. Toxicological Sciences, 165(1), 198–212. https://doi.org/10.1093/toxsci/kfy152
17. Luechtefeld, T., Rowlands, C., & Hartung, T. (2018). Big-data and machine learning to revamp computational toxicology and its use in risk assessment. Toxicology Research, 7(5), 732–744. https://doi.org/10.1039/c8tx00051d
18. MacArthur Clark, J. (2017). The 3Rs in research: a contemporary approach to replacement, reduction and refinement. British Journal of Nutrition, 120(s1), S1–S7. https://doi.org/10.1017/s0007114517002227
19. Maesh, K. and Bugusu, B. (2007). Food packaging--roles, materials, and environmental issues. J Food Sci. 72(3):R39-55. https://www.ncbi.nlm.nih.gov/pubmed/17995809
20. Malloy, T., & Beryt, E. (2016). Leveraging the new predictive toxicology paradigm: alternative testing strategies in regulatory decision-making. Environmental Science: Nano, 3(6), 1380–1395. https://doi.org/10.1039/c6en00202a
21. Meigs, L. (2018). Animal testing and its alternatives – the most important omics is economics. ALTEX, 275–305. https://doi.org/10.14573/altex.1807041
22. Mitchell, C. (2014). Global Regulation of Food Additives. https://pubs.acs.org/doi/pdf/10.1021/bk-2014-1162.ch001
23. Mitchell, R. E., Fraser, A. M., & Bearon, L. B. (2007). Preventing food-borne illness in food service establishments: Broadening the framework for intervention and research on safe food handling behaviors. International Journal of Environmental Health Reseach, 17(1), 9-24. http://doi.org/10.1080/09603120601124371
24. Mordor Intelligence. (2019). Food Additives Market – growth, trends, and forecast (2020-2025). https://www.mordorintelligence.com/industry-reports/global-food-additives-market-industry
25. National Research Council (US) Committee on Diet, Nutrition, and Cancer. Washington (DC): National Academies Press (US); 1983.
26. Neltner, T. G., Alger, H. M., Leonard, J. E., & Maffini, M. V. (2013). Data gaps in toxicity testing of chemicals allowed in food in the United States. Reproductive Toxicology, 42, 85–94. https://doi.org/10.1016/j.reprotox.2013.07.023
27. Patlewicz, G., Ball, N., Becker, R. A., Booth, E. D., Cronin, M. T. D., Kroese, D., Steup, D., van Ravenzwaay, B. and Hartung, T. (2014). Read-across approaches - misconceptions, promises and challenges ahead. ALTEX, 31(4), 387–396. https://doi.org/10.14573/altex.1410071
28. Processing Aids Used in Modern Food Production. International Food Information Council. 23 May 2014. Archived from the original on 19 October 2016.
29. Roca-Saavedra, P., Mendez-Vilabrille, V., Miranda, J. M., Nebot, C., Cardelle-Cobas, A., Franco, C. M., & Cepeda, A. (2017). Food additives, contaminants and other minor components: effects on human gut microbiota—a review. Journal of Physiology and Biochemistry, 74(1), 69–83. https://doi.org/10.1007/s13105-017-0564-2